Cellular Physiology and Biochemistry

Cell Physiol Biochem 2018;50:1429-1440
DOI: 10.1159/000494605
Published online: 25 October 2018
Accepted: 17 October 2018







**Original Paper**

# Circulating Long Noncoding RNAs as Biomarkers for Predicting Head and Neck Squamous Cell Carcinoma

Yao Yao[a] Xinyuan Chen[a] Shuai Lu[b,c,d] Chao Zhou[b,c,d] Guolong Xu[b,c,d] Zhongyi Yan[b,e,f] Jianrong Yang[e,f] Tonghua Yu[a] Wei Chen[a] Yichun Qian[a] Shounian Ding[g] Junwei Tang[h] Yun Chen[a,b,c,d] Yuan Zhang[a]

[a]Department of Head and Neck Surgery, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & Nanjing Medical University Affiliated Cancer Hospital, Nanjing, [b]Department of Immunology, Nanjing Medical University, Nanjing, [c]Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, [d]Key Laboratory of Human Functional Genomics of Jiangsu Province, Jiangsu Diabetes Center, Nanjing Medical University, Nanjing, [e]Jiangsu Key Laboratory of Oral Disease, Nanjing Medical University, Jiangsu, [f]Department of Oral and Maxillofacial Surgery, Affiliated Hospital of Stomatology, Nanjing Medical University, Jiangsu, [g]School of Chemistry and Chemical Engineering, Southeast University, Nanjing, [h]Department of Colorectal Surgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

**Key Words**
Plasma • LncRNA • Microarrays • Risk score function • ROC curve

**Abstract**
***Background/Aims:*** The anatomical complexity of the head and neck region and the lack of sufficiently specific and sensitive biomarkers often lead to the diagnosis of head and neck squamous cell carcinoma (HNSCC) at advanced stages. To identify novel biomarkers for early diagnosis of primary HNSCC through a minimally invasive method, we investigated circulating long noncoding RNA (lncRNA) levels in plasma of HNSCC patients. ***Methods:*** The global lncRNA expression profiles of HNSCC patients were measured using microarray and next-generation RNA-sequencing (RNA-seq) data from both circulating and tissue samples. The diagnosis prediction model based on the lncRNA signatures and clinical features was evaluated by multi-stage validation and risk score analysis. ***Results:*** The data showed that 432 lncRNA transcripts were differentially expressed by fold changes of >4 in circulating samples and 333 in tissues samples, respectively. Only 12 lncRNAs consistently emerged in these two kinds of samples. After the risk score analysis including a multistage validation, we identified three lncRNAs, namely, HOXA11-AS, LINC00964 and MALAT1, which were up-regulated in the plasma of HNSCC patients compared with those in healthy controls with merged areas under

Y. Yao and X. Chen contributed equally to this work.

Yuan Zhang
and Yun Chen

Department of Head and Neck Surgery, Jiangsu Cancer Hospital
Nanjing, 210009 (China)
E-Mail yuan_zhang63@163.com; chenyun@njmu.edu.cn

the curve (AUCs) in training and validation sets of 0.925 and 0.839, respectively. ***Conclusion:*** HOXA11-AS, LINC00964 and MALAT1 might be potential circulating biomarkers for the early detection of HNSCC in the future.



## Introduction

Head and neck cancer (HNC) accounts for more than 630, 000 newly diagnosed cancer cases and over130, 000 cancer deaths annually [1, 2]. Head and neck squamous cell carcinoma (HNSCC), the most common malignant lesion in the head and neck region, originates in the oral cavity, oropharynx, hypopharynx, and larynx [3]. Although significant advances have been made in comprehensive therapeutics, the 5-year overall survival rate for patients with HNSCC has remained at approximately 50% in recent decades, mainly due to local recurrence and regional lymph node and distant metastasis [4, 5]. The anatomical complexity of the head and neck region and the lack of sufficiently specific and sensitive biomarkers often lead to the diagnosis of HNSCCs at advanced stages, which results in poor prognosis and survival rates [6]. Therefore, the identification of novel biomarkers for the early diagnosis of primary HNSCC will play an essential role in improving outcomes [4, 7].

Long non-coding RNAs (lncRNAs) are a newly-identified class of RNAs that are more than 200 nucleotides in length without protein-coding functions [8]. Currently, high-throughput microarray profiles and RNA sequencing (RNA-seq) experiments have identified a large number of abnormally expressed lncRNAs in various cancers [9-14]. The expression of lncRNAs is generally tissue-specific [15, 16]. Similar to proteins and microRNAs, lncRNAs have been reported as novel diagnostic and prognostic biomarkers in cancer not only in tumor tissues but also in blood [17-21]**.** In HNSCC, HOTAIR, UCA-1 and GAS5 are dysregulated and associated with poor prognosis [22-24]. Most recently, a study focused on plasma testing has shown that circulating lncRNAs RP11-160H22.5, XLOC_014172 and LOC149086 might be potential biomarkers for tumorigenesis and the latter two might be potential biomarkers for metastasis in hepatocellular carcinoma [25]. Similarly, circulating GAS5 levels may serve as a potential biomarker in predicting treatment response in HNSCC patients treated with chemoradiotherapy [26].

To the best of our knowledge, although the biomarkers of HNSCC have been extensively investigated, studies focused on the diagnostic effects of circulating lncRNAs of this disease are rare. In the present study, we investigated the potential use of circulating lncRNAs in plasma as biomarkers of HNSCC. First, we identified aberrant expression of lncRNAs by analyzing the microarray and RNA-seq data from HNSCC patients and cancer-free control individuals (both circulating and tissue samples). Then, we employed reverse-transcription quantitative polymerase chain reaction (RT-qPCR) to confirm and validate the results in plasma from more HNSCC patients and cancer-free controls. Therefore, we generated a panel of differentially expressed lncRNAs in plasma that might serve as novel biomarkers for the diagnosis of HNSCC.

## Materials and Methods

*Study design*

A total of 100 patients with HNSCC who were referred to the Department of Head and Neck Surgery at Nanjing Medical University Affiliated Cancer Hospital from 2012 to 2015 and 100 matched healthy individuals without any apparent disease were enrolled in the study. This study was performed in accordance with the Declaration of Helsinki (fifth revision, October 2000) of the World Medical Association and approved by the Ethical Committee of Nanjing Medical University Affiliated Cancer Hospital. Written informed consent was provided by all of the study participants. None of the patients had received any other treatment, such as radiotherapy or chemotherapy, prior to surgery. The diagnosis of HNSCC was confirmed by histopathological analysis of tumor tissue from the surgical resection specimen. Each tissue specimen was stored in liquid

nitrogen and blood samples were stored at -80°C until use. The corresponding tissues adjacent to tumors were defined as adjacent normal tissues (located >3 cm away from the tumor) and confirmed pathologically. Blood samples collected from each donor were placed in EDTA-containing anticoagulant tubes. Plasma was separated by centrifugation at 800 g for 10 min at room temperature, followed by a 15-min high-speed centrifugation at 10, 000 g at room temperature for complete removal of cell debris. The plasma supernatant was recovered and stored at -80°C until analysis. The histological types of tumors were classified according to the criteria determined by the World Health Organization. The tenth edition of American Joint Committee on Cancer (AJCC) tumor-node-metastasis (TNM) systems was used for staging the disease based on clinical, radiological, and pathological findings. The detailed information of patients and healthy controls is summarized in Table 1.

**Table 1.** Clinicopathological characteristics of HNSCC and cancer-free control samples. [a] Student's t-test. [b] Chi-square test

| Characteristics | HNSCC | Control | P value |
|---|---|---|---|
| N | 100 | 100 | |
| Mean age (SE), years | 54.33(8.12) | 56.11(7.58) | 0.22[a] |
| Sex (male/female) | 61/39 | 59/41 | 0.77[b] |
| Tumor site | | | |
| Oral cavity/Oropharynx | 55 | | |
| Larynx/Hypopharynx | 45 | | |
| Histological grade | | | |
| G1 | 41 | | |
| G2 | 37 | | |
| G3 | 22 | | |
| TNM stage(I:II:III) | 43:31:26 | | |
| Lymph node metastasis | | | |
| Yes | 67 | | |
| No | 33 | | |

*Screening phase*

The screening phase was divided into a training set and validation set. Twenty paired samples including samples from 20 patients with HNSCC and 20 healthy controls were used in the training set, while 100 samples from patients diagnosed with HNSCC and 100 samples from healthy controls served as the validation set.

*Training set*

All candidate lncRNAs were tested in a group of 20 plasma and tissue samples obtained from patients, as well as 20 plasma from healthy controls. The tissue samples included paired tumor tissues and adjacent normal tissues. All patients were pathologically diagnosed with HNSCC. The expression levels of candidate lncRNAs were analyzed in all samples, and the comparative $2^{-\Delta\Delta Ct}$ method was used to analyze the difference between patients and healthy controls.

*Validation set*

A case-control study was designed to validate the obvious difference of relative expression levels of the selected potential biomarkers in a cohort of 100 patients and 100 healthy controls.

*RNA extraction*

Total RNA was extracted from 50 mg tissue and 300 μL plasma via TRIzol reagent according to the manufacturer's protocol (Invitrogen, Carlsbad, CA). Total RNA from each sample was quantified using NanoDrop ND-1000 (Thermo Fisher Scientific Inc., Waltham, MAc).

*LncRNA microarray and data analysis*

RNA extracted from plasma samples from three patients diagnosed with HNSCC and three healthy controls, regarded as circulating samples, was selected for lncRNA microarray analysis using the Agilent (California, USA) microarray platform. Sample preparation and microarray hybridization were performed by CapitalBio Corporation (Beijing, China) according to the manufacturer's standard protocols as previously described [27]. Briefly, cDNA labeled with a fluorescent dye (Cy5 and Cy3-dCTP) was synthesized and hybridized to the Agilent human lncRNA+mRNA Array V4.0 (Agilent, Santa Clara, CA, United States). The data was analyzed for summarization, normalization and quality control by using the Agilent GeneSpring software V13.0 (Agilent Technologies, Inc). The data were log2 transformed and median centered by genes using the Adjust Data function of CLUSTER 3.0 software (University of Tokyo, Human Genome Center, Tokyo, Japan) and further analyzed with hierarchical clustering with average linkage. Tree visualization were performed by Java TreeView (Stanford University School of Medicine).

*Illumina HiSeq 2500 sequencing and data analysis*

RNA extracted from three HNSCC patients after surgery including paired tumor tissues and adjacent normal tissues, regarded as tissue samples, was subjected to RNA-seq analysis on the Illumina Hiseq 2500 platform (Illumina, USA) by CapitalBio Corporation according to the manufacturer's instructions described previously [28]. In brief, sequencing libraries were generated using the NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina® (NEB, USA). The libraries were sequenced using high-throughput 125-bp paired-end mode by the Illumina Hiseq 2500 device. For the data analysis, base calls were performed using bcl2fastq2 (v2.20.0). Reads were filtered using NGSQC-toolkit (v2.3.3), and aligned to the genome using the split read aligner TopHat (v2.0.13) with the mismatches set at 2 and with the default parameters. The software package limma (v.10) was used to calculate the fold-change of transcripts and to screen all differentially expressed genes. Heatmap and cluster analyses were applied to the expression data set.

*RT-qPCR*

RT-qPCR was used to validate the data from the microarray and RNA-seq results. Total RNA from tissue and plasma samples was reverse transcribed using the reverse transcription kit (Takara, Tokyo, Japan). RT-qPCR was then performed using an ABI Prism 7900HT system (Applied Biosystems, CA, USA).

*Risk score analysis*

Risk score analysis was performed to evaluate the associations between tissue and plasma lncRNA expression levels as previously described [25]. The upper 95% reference interval of each lncRNA in the control group was set as the threshold to code the expression level of the corresponding lncRNA for each sample as 0 and 1 in the training set. A risk score function (RSF) to predict the HNSCC group was defined according to a linear combination of the expression level for each lncRNA. For example, the RSF for sample i using information from three lncRNAs was: rsfi=$\sum$3j-1Wj.sij, where, sij is the risk score for lncRNA j in sample i, and Wj is the weight of the risk score of lncRNA j. The risk score of three lncRNAs was calculated using the weight of the regression coefficient that was derived from the univariate logistic regression analysis of each lncRNA. Samples were ranked according to their RSF and then divided into a high-risk group, representing the HNSCC patients, and a low-risk group, representing the predicted control individuals. Frequency tables and receiver operating characteristic (ROC) curves were then used to evaluate the diagnostic effects of the profiling, find the appropriate cut-off point and validate the procedure and cutoffs in the next validation set.

*Statistical analysis*

Data are presented as the mean (SEM). Student's t-test and chi-square tests of variance were used to evaluate clinicopathological characteristics. The paired and unpaired student's t-test were used to compare significant differences in tissue and plasma lncRNA expression between patients and controls. Analysis of area under the ROC curve (AUC) was used to estimate the effectiveness of lncRNAs for prediction. Statistical analysis was performed using SPSS 22.0 software and GraphPad Prism 6.0. In all cases, $p<0.05$ was considered statistically significant. All p values were two-sided.

## Results

*High-throughput microarray and RNA-seq detection of circulating and tissue lncRNAs*

The Agilent microarray and the Illumina HiSeq 2500 platform was employed to detect the lncRNAs derived from circulating and tissue samples. We randomly chose plasma samples obtained from three HNSCC patients before operation and from three healthy controls. The tissue samples included three matched primary HNSCC tissues and adjacent nontumor mucous membrane tissues. As presented in Fig. 1A-1D, aberrant expression of lncRNAs was detected in both circulating and tissue samples. To screen the candidate biomarkers for predicting HNSCC, we filtered the differentially expressed lncRNAs with the following parameters: i, lncRNAs with > 80% expression; ii, lncRNAs with fold changes > 4; and iii, lncRNAs increased in both circulating and tissue samples. As shown in Fig. 1E, 432 aberrant lncRNAs increased in circulating samples from HNSCC patients, while 333 lncRNAs

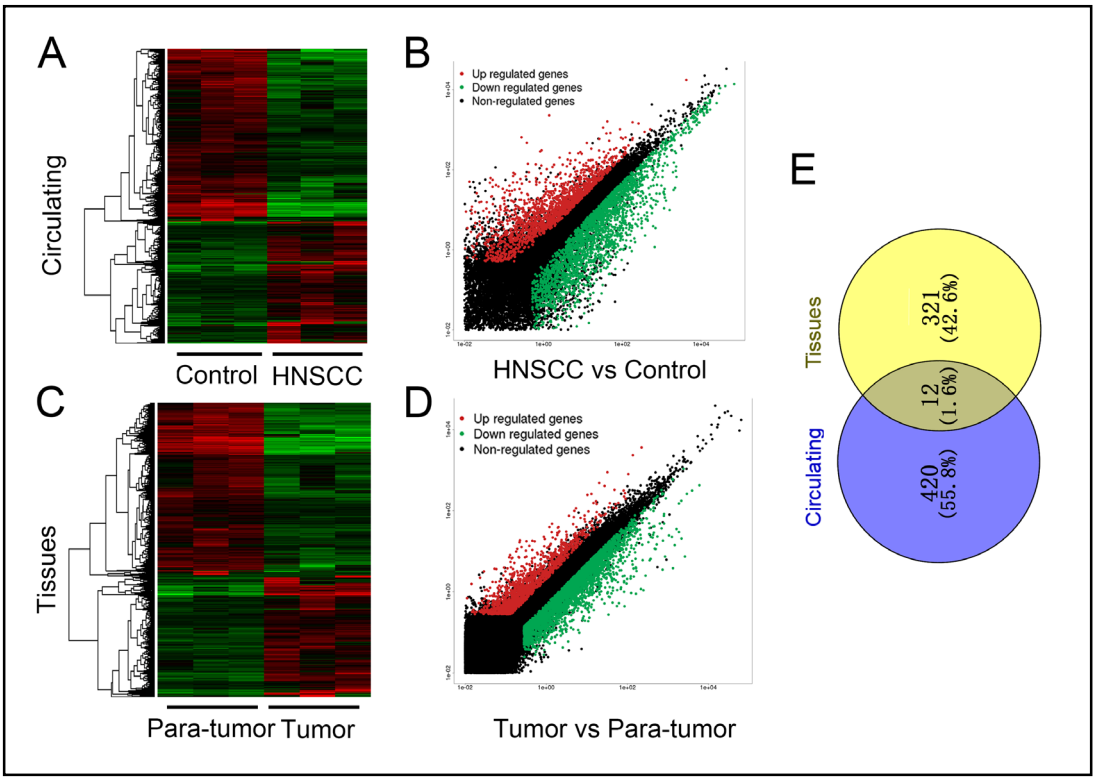


**Fig. 1.** lncRNA expression profiling in circulating and tissue samples. A: Cluster analysis and scatter plot of the differentially expressed lncRNAs (plasma samples from three patients with HNSCC and three cancer-free controls). B: Cluster analysis and scatter plot of the differentially expressed lncRNAs (three paired tumor tissues and adjacent normal tissues). C: Venny analysis of upregulated lncRNAs in both circulating and tissue samples.

increased in tissue samples (tumor tissues). Only 12 lncRNAs were consistent with all the parameters mentioned above and were was further validated in the subsequent screening phase.

*Validation of significantly dysregulated lncRNAs in plasma by RT-qPCR*

To verify the microarray data in plasma, we validated the expression level of lncRNAs by RT-qPCR in plasma from 20 cancer patients and 20 cancer-free controls. No significant differences were detected in the distribution of age or sex between the cancer patients and cancer-free controls. The results showed that among the 12 candidate lncRNAs, six lncRNAs were significantly different in circulating samples. (Fig. 2).

Next, the 6 candidate lncRNAs were confirmed in 20 tissue samples, including paired tumor and adjacent normal tissues, and 20 circulating samples from HNSCC patients and cancer-free patients. The 20 paired samples from the same patients are described in the previous paragraph. Only three lncRNAs were confirmed to be increased in both HNSCC patient plasma samples and tumor tissues (Fig. 3).

Finally, we validated the 3 aberrantly expressed lncRNAs in samples from another 100 patients with HNSCC and 100 cancer-free controls. As presented in Fig. 4, 3 out of 6 lncRNAs including HOXA11-AS, LINC00964 and metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) were confirmed to be significant.

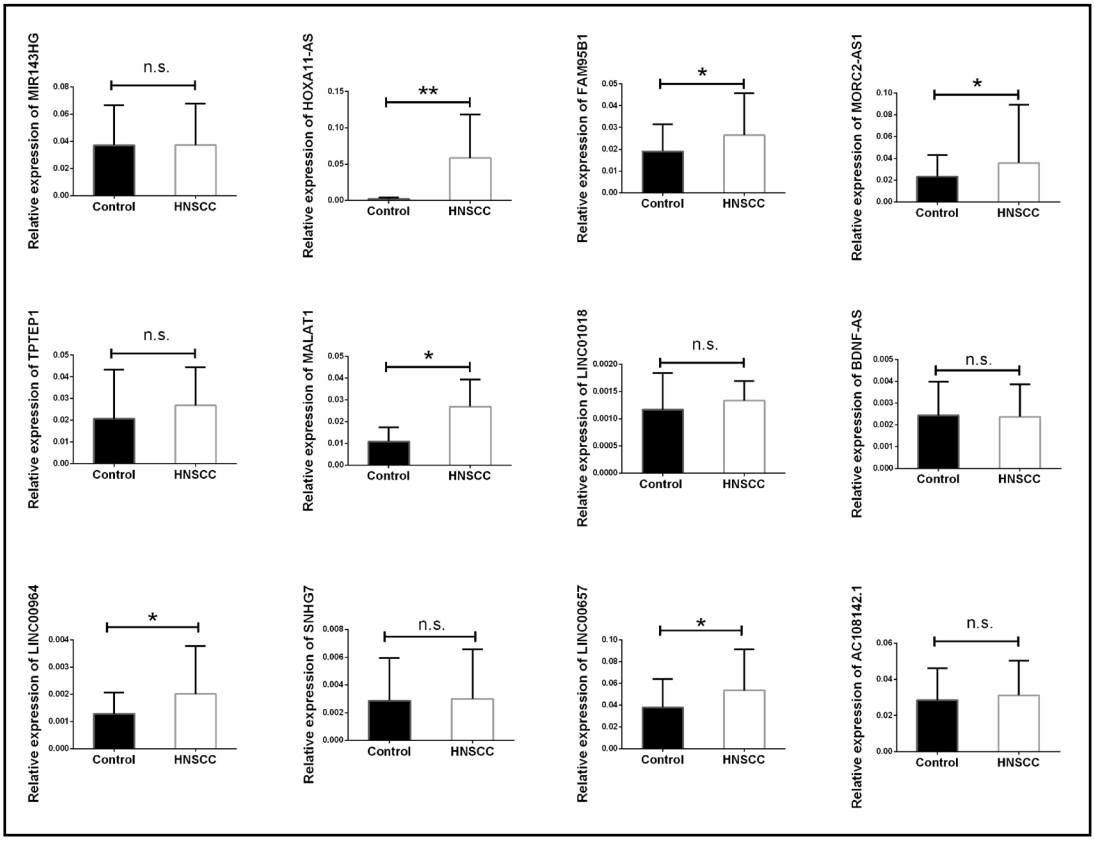


**Fig. 2.** Validation of candidate lncRNAs in training set. Paired plasma samples from 20 HNSCC patients, and 20 cancer-free controls were used for RT-qPCR analysis. Data are presented as the mean ± SEM. Data were analyzed with Student's t-test. n.s. indicates no significance, * indicates $p<0.05$ and ** indicates $p<0.01$.

*Risk score analysis*

Further experiments were conducted to explore the accuracy and specificity of these 3 lncRNAs as potential HNSCC biomarkers. Throughout the multiphase testing and analysis, a profile of three lncRNAs might be considered a potential signature for the diagnosis of HNSCC.

Risk score analysis was applied to evaluate the diagnostic value of the three lncRNAs. First, we divided the control group and case group in the training set according to the upper 95% credibility interval (95% CI) in the control group. The risk score was calculated based on the results of logistic regression analysis. All plasma samples were then divided into a high-risk group, indicating the possible HNSCC patients, and a low-risk group, representing the predicted controls. We defined the cutoff value as the maximal value of sensitivity + specificity. The positive predictive value (PPV) and negative predictive value (NPV) were calculated as 70% and 80% in the training set, respectively. We further applied the same value to calculate the risk score of samples in the validation sets, and, the PPV and NPV were 88% and 81%, respectively (Table 2).

ROC curve analysis was used to evaluate the value of lncRNAs for predicting HNSCC. The AUCs of the three validated lncRNA were 0.773, 0.387 and 0.918, respectively, while the combination of the three lncRNAs displayed a moderate ability of discriminating between HNSCC patients and controls with an AUC of 0.925 (Fig. 5A).

Next, ROC curve analysis was used to evaluate the value of lncRNAs for predicting HNSCC in the validation set. The AUCs of the three validated lncRNAs were 0.659, 0.687 and 0.718, respectively, while the combination of the three factors displayed a moderate ability of discriminating between HNSCC patients and controls with an AUC of 0.839 (Fig. 5B).

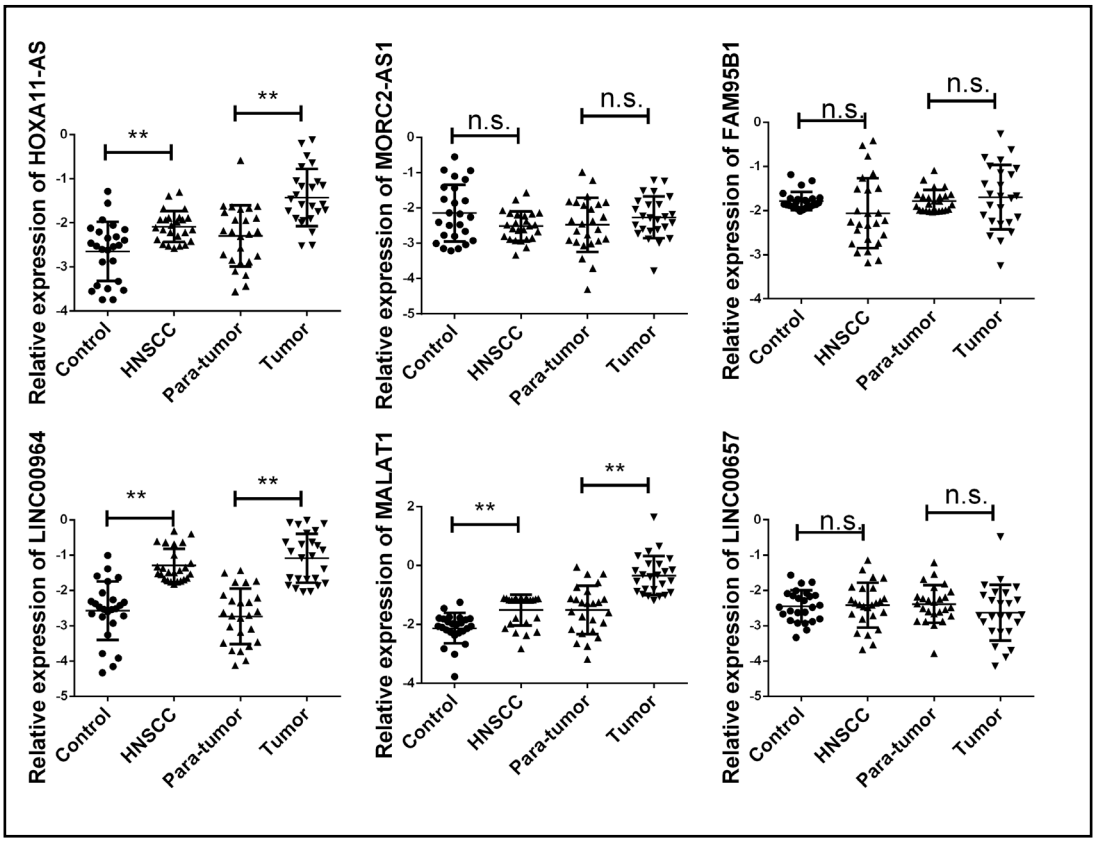


**Fig. 3.** Validation of candidate lncRNAs in both circulating and tumor samples. Circulating samples including paired plasma samples from 20 HNSCC patients and 20 cancer-free controls and tissue samples including 20 paired tumor tissues and adjacent normal tissues were used for RT-qPCR analysis. Data are presented as the mean ± SEM. Data were analyzed with Student's t-test. n.s. indicates no significance and ** indicates p<0.01.

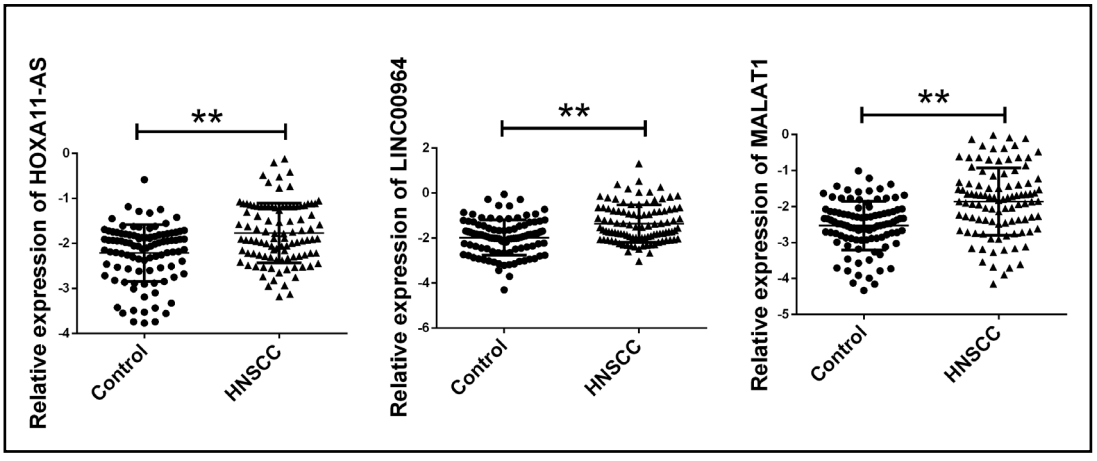


**Fig. 4.** Validation of candidate lncRNAs in the validation set. Paired plasma samples from 100 HNSCC patients and 100 cancer-free controls were used for RT-qPCR analysis. Data are presented as the mean ± SEM. Data were analyzed with Student's t-test, ** indicates p<0.01.

## Discussion

Although surprising advances have been made in cancer research and therapy, the survival of patients with HNSCC has not significantly improved in the last few decades [4, 5, 29]. There remains to date a lack of effective tumor-specific biomarkers for the early detection and prognostic prediction of HNSCC. Therefore, developing new, highly sensitive and specific biomarkers for the diagnosis of HNSCC is urgently needed to improve patient outcomes. Previously published studies on biomarkers in HNSCC have mainly focused on proteins and microRNAs. Epidermal growth factor receptor (EGFR), for instance, is associated with poor prognosis in HNSCC and can serve as a predictive biomarker for therapy response [30-32]. MiR-21 expression in the tumor stroma has a negative prognostic value in oral squamous cell carcinoma [33]. Low levels of both miR-205 and let-7d expression in HNSCC tumors are significantly associated with poor survival [34].

Recently, lncRNAs have emerged as new regulators of cancer progression [14, 35]. MALAT1 is overexpressed and has been

**Table 2.** Risk score analysis in HNSCC and cancer-free control plasma samples. [a] PPV, positive predictive value. [b] NPV, negative predictive value

| Score | 0-6.89 | 6.89-12.11 | PPV[a] | NPV[b] |
|---|---|---|---|---|
| Training set | | | 0.70 | 0.80 |
| HNSCC | 6 | 14 | | |
| Control | 16 | 4 | | |
| Validation set | | | 0.88 | 0.81 |
| HNSCC | 12 | 88 | | |
| Control | 81 | 19 | | |

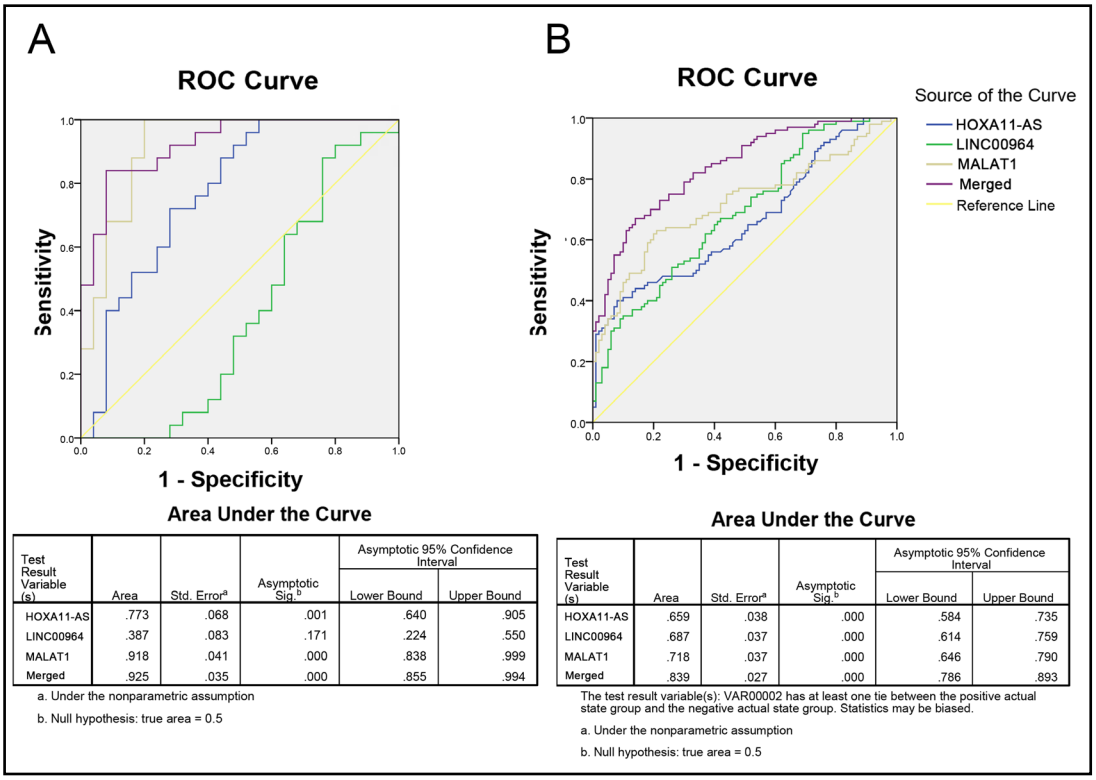


**Fig. 5.** ROC analysis of the three potential biomarkers for HNSCC using risk score analysis. A: ROC curve analysis was conducted for discriminating between HNSCC patients and controls using the three-lncRNA signature. ROC curve analysis was performed for the three-lncRNA signature to separate 20 HNSCC patients from 20 controls in the training set with the AUC presented in the lower panel. Merged value indicates the combination of HOXA11-AS, LINC00964 and MALAT1. B: ROC curve for the three-lncRNA signature to separate 100 HNSCC patients from 100 controls in the validation set with the AUC presented in the lower panel. Merged value indicates the combination of HOXA11-AS, LINC00964 and MALAT1.

used as a new prognostic factor of survival in multiple cancers [36, 37]. MALAT1 may be a promising therapeutic target for suppressing cancer progression and drug resistance [38-40]. Low expression levels of both miR-26a and MEG3 are an independent prognostic factor for poor clinical outcomes in tongue squamous cell carcinoma (TSCC) patients [41]. LINC00152 derived from the Gene Expression Omnibus (GEO) database might serve as a potential biomarker for the early detection and prognostic prediction of TSCC [42]. A three-lncRNA panel (KTN1-AS1, LINC00460 and RP5-894A10.6) derived from the Atlas of Noncoding RNA in Cancer (TANRIC) database might be a novel biomarker for the accurate prognostic prediction of patients with HNSCC [43]. Circulating lncRNAs in peripheral blood can also be excellent diagnostic candidate biomarkers because of their stability in circulating and ability to be detected by minimally invasive methods in many cancers, including HNSCC [44-47].

To the best of our knowledge, although studies on the effects of lncRNAs on HNSCC are still ongoing, the global profiling of dysregulated lncRNAs in the plasma of HNSCC patients has never been explored utilizing both microarray profiling and next-generation sequencing. This is the first study assessing universal circulating lncRNAs for HNSCC diagnosis according to microarray and RNA-seq data obtained from both blood and tissues. In this study, regardless of the various tumor sites, we attempted to discover universal circulating lncRNA biomarkers of HNSCC.

Using the microarray and RNA-seq method, we could quantify the dysregulated expression levels of HNSCC-associated lncRNAs derived from both plasma and tissue samples. The data showed that 432 lncRNA transcripts were significantly differentially expressed by fold changes of > 4 in circulating samples and 333 in tissues samples, respectively. Only 12 lncRNAs consistently emerged in these two kinds of samples.

Then, utilizing risk score analysis involving a multistage validation, we evaluated the association between the lncRNA expression levels in plasma samples and the clinical features of HNSCC patients including age, sex, histologic grade, TNM stage, lymph node and distant metastasis. At this stage, we screened and discovered three novel candidate circulating lncRNAs, namely, HOXA11-AS, LINC00964 and MALAT1, which act as biomarkers in the early diagnosis of primary HNSCC. We selected these three candidate lncRNAs because they have not only been demonstrated to be significantly dysregulated in the plasma of HNSCC patients, but also been found to be stable in circulation when subjected to harsh conditions. The optimal cutoff values from ROC curve analyses indicated that the combined detection of these three biomarkers could discriminate HNSCC patients from healthy controls with a relatively high sensitivity (88%) and specificity (81%) and was more accurate than the individual detection of each biomarker. Furthermore, the ROC analysis showed that the combination of lncRNAs HOXA11-AS, LINC00964 and MALAT1 is suitable for the early diagnosis of HNSCC with merged AUCs in the training and validation sets of 0.925 and 0.839, respectively. From the ROC analysis, we found a remarkable difference in LINC00964 levels between the training set and validation set. Since the samples for the training set were chosen at random, this might have resulted in a poor ability to distinguish HNSCC patients from controls; however, the predictive ability can improve with an increase in the number of samples. Thus, the results might be more reliable with larger sample sizes.

HNCs encompass malignancies originating from multiple anatomical sites. Despite the anatomical diversity of the head and neck region, more than 90% of HNC cases are classified as HNSCC. The clinical management of HNSCC involves several challenges including the early detection of primary tumors.

HNSCC represents a genetically heterogeneous group of tumors in which the molecular characteristics and clinical outcomes vary widely. In this study, we focused on novel potential biomarkers related to HNSCC. For example, one of the most well-studied biomarkers of HNSCC is EGFR, which is commonly overexpressed in HNSCC and is a negative prognostic factor associated with poor local control and survival. Despite significant differences at the molecular level among certain HNSCC subtypes, we attempted to discover circulating lncRNA biomarkers for HNSCC in general.

In this study, we performed simultaneous analysis of tissue and plasma samples, which is important for extracting some information on common variations from data sets. However, the results need to be sufficiently validated in the future for a comprehensive understanding of molecular deregulations in HNSCC cells *in vitro* and *in vivo*.

## Conclusion

In conclusion, in this study, we identified three lncRNAs, namely, HOXA11-AS, LINC00964 and MALAT1, as potential biomarkers of tumorigenesis. However, this is only a preliminary study, and is limited by a small sample size. A deeper understanding of the potential function of the three lncRNAs in the regulation of HNSCC pathogenesis is necessary for us in the future.

## Abbreviations

HNSCC (head and neck squamous cell carcinoma); lncRNAs (long noncoding RNAs); RT-qPCR (reverse transcription quantitative polymerase chain reaction); ROC (curve, receiver operating characteristic curve); AUC (area under the curve); RSF (risk score function).

## Acknowledgements

This work was funded by Jiangsu Provincial key research development program (BE2016796, BE2018750), the Six Talent Peaks Project (2016-WSW-022), the project of Nanjing Science and Technology Commission (201611006), the China Postdoctoral Science Foundation(2017M621676), National Natural Science Foundation (81702338) and the Basic Research Project of Jiangsu Province (Natural Science Foundation) (BK20171090).

## Disclosure Statement

The authors declare that they have no financial conflicts of interest.

## References

1 Siegel RL, Miller KD, Jemal A: Cancer Statistics, 2017. CA Cancer J Clin 2017;67:7-30.
2 Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ, He J: Cancer statistics in China, 2015. CA Cancer J Clin 2016;66:115-132.
3 Marur S, Forastiere AA: Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment. Mayo Clin Proc 2016;91:386-396.
4 Agrawal N, Frederick MJ, Pickering CR, Bettegowda C, Chang K, Li RJ, Fakhry C, Xie TX, Zhang J, Wang J, Zhang N, El-Naggar AK, Jasser SA, Weinstein JN, Trevino L, Drummond JA, Muzny DM, Wu Y, Wood LD, Hruban RH, Westra WH, Koch WM, Califano JA, Gibbs RA, Sidransky D, Vogelstein B, Velculescu VE, Papadopoulos N, Wheeler DA, Kinzler KW, Myers JN: Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science 2011;333:1154-1157.
5 Rahimy E, Kuo SZ, Ongkeko WM: Evaluation of Non-Coding RNAs as Potential Targets in Head and Neck Squamous Cell Carcinoma Cancer Stem Cells. Curr Drug Targets 2014;15:1247-1260.
6 Kowalski LP, Carvalho AL: Influence of time delay and clinical upstaging in the prognosis of head and neck cancer. Oral Oncol 2001;37:94-98.
7 Dunne S, Mooney O, Coffey L, Sharp L, Desmond D, Timon C, O'Sullivan E, Gallagher P: Psychological variables associated with quality of life following primary treatment for head and neck cancer: a systematic review of the literature from 2004 to 2015. Psycho-Oncology 2017;26:149-160.

8 Ponting CP, Oliver PL, Reik W: Evolution and functions of long noncoding RNAs. Cell 2009;136:629-641.
9 Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, Sukumar S, Chang HY: Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 2010;464:1071-1076.
10 Hu L, Wu Y, Tan D, Meng H, Wang K, Bai Y, Yang K: Up-regulation of long noncoding RNA MALAT1 contributes to proliferation and metastasis in esophageal squamous cell carcinoma. J Exp Clin Canc Res 2015;34:7.
11 Yang F, Huo XS, Yuan SX, Zhang L, Zhou WP, Wang F, Sun SH: Repression of the Long Noncoding RNA-LET by Histone Deacetylase 3 Contributes to Hypoxia-Mediated Metastasis. Mol Cell 2013;49:1083-1096.
12 Hirata H, Hinoda Y, Shahryari V, Deng GR, Nakajima K, Tabatabai ZL, Ishii N, Dahiya R: Long Noncoding RNA MALAT1 Promotes Aggressive Renal Cell Carcinoma through Ezh2 and Interacts with miR-205. Cancer Res 2015;75:1322-1331.
13 Liu Q, Huang JG, Zhou NJ, Zhang ZQ, Zhang AL, Lu ZH, Wu FT, Mo YY: LncRNA loc285194 is a p53-regulated tumor suppressor. Nucleic Acids Res 2013;41:4976-4987.
14 Wang L, Yang F, Jia LT, Yang AG: Missing Links in Epithelial-Mesenchymal Transition: Long Non-Coding RNAs Enter the Arena. Cell Physiol Biochem 2017;44:1665-1680.
15 Gibb EA, Brown CJ, Lam WL: The functional role of long non-coding RNA in human carcinomas. Mol Cancer 2011;10:38.
16 Schmitz SU, Grote P, Herrmann BG: Mechanisms of long noncoding RNA function in development and disease. Cell Mol Life Sci 2016;73:2491-2509.
17 Akrami R, Jacobsen A, Hoell J, Schultz N, Sander C, Larsson E: Comprehensive analysis of long non-coding RNAs in ovarian cancer reveals global patterns and targeted DNA amplification. PLoS One 2013;8:e80306.
18 Li J, Chen Z, Tian L, Zhou C, He MY, Gao Y, Wang S, Zhou F, Shi S, Feng X, Sun N, Liu Z, Skogerboe G, Dong J, Yao R, Zhao Y, Sun J, Zhang B, Yu Y, Shi X, Luo M, Shao K, Li N, Qiu B, Tan F, Chen R, He J: LncRNA profile study reveals a three-lncRNA signature associated with the survival of patients with oesophageal squamous cell carcinoma. Gut 2014;63:1700-1710.
19 Fatima R, Akhade VS, Pal D, Rao SM: Long noncoding RNAs in development and cancer: potential biomarkers and therapeutic targets. Mol Cell Ther 2015;3:5.
20 Zhou M, Guo M, He D, Wang X, Cui Y, Yang H, Hao D, Sun J: A potential signature of eight long non-coding RNAs predicts survival in patients with non-small cell lung cancer. J Transl Med 2015;13:231.
21 Zhou M, Zhao H, Xu W, Bao S, Cheng L, Sun J: Discovery and validation of immune-associated long non-coding RNA biomarkers associated with clinically molecular subtype and prognosis in diffuse large B cell lymphoma. Mol Cancer 2017;16:16.
22 Fang Z, Wu L, Wang L, Yang Y, Meng Y, Yang H: Increased expression of the long non-coding RNA UCA1 in tongue squamous cell carcinomas: a possible correlation with cancer metastasis. Oral Surg Oral Med Oral Pathol Oral Radiol 2014;117:89-95.
23 Li DD, Feng JP, Wu TY, Wang YD, Sun YN, Ren JY, Liu M: Long Intergenic Noncoding RNA HOTAIR Is Overexpressed and Regulates PTEN Methylation in Laryngeal Squamous Cell Carcinoma. Am J Pathol 2013;182:64-70.
24 Gee HE, Buffa FM, Camps C, Ramachandran A, Leek R, Taylor M, Patil M, Sheldon H, Betts G, Homer J, West C, Ragoussis J, Harris AL: The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis. Br J Cancer 2011;104:1168-1177.
25 Tang JW, Jiang RQ, Deng L, Zhang XD, Wang K, Sun BC: Circulation long non-coding RNAs act as biomarkers for predicting tumorigenesis and metastasis in hepatocellular carcinoma. Oncotarget 2015;6:4505-4515.
26 Fayda M, Isin M, Tambas M, Guveli M, Meral R, Altun M, Sahin D, Ozkan G, Sanli Y, Isin H, Ozgur E, Gezer U: Do circulating long non-coding RNAs (lncRNAs) (LincRNA-p21, GAS 5, HOTAIR) predict the treatment response in patients with head and neck cancer treated with chemoradiotherapy? Tumor Biol 2016;37:3969-3978.
27 Lin C, Zhang S, Wang Y, Wang Y, Nice E, Guo C, Zhang E, Yu L, Li M, Liu C, Hu L, Hao J, Qi W, Xu H: Functional Role of a Novel Long Noncoding RNA TTN-AS1 in Esophageal Squamous Cell Carcinoma Progression and Metastasis. Clin Cancer Res 2018;24:486-498.
28 Shi Y, Ye P, Long X: Differential Expression Profiles of the Transcriptome in Breast Cancer Cell Lines Revealed by Next Generation Sequencing. Cell Physiol Biochem 2017;44:804-816.

29 Rivera C: Essentials of oral cancer. Int J Clin Exp Pathol 2015;8:11884-11894.
30 Shin DM, Ro JY, Hong WK, Hittelman WN: Dysregulation of epidermal growth factor receptor expression in premalignant lesions during head and neck tumorigenesis. Cancer Res 1994;54:3153-3159.
31 Rubin Grandis J, Melhem MF, Gooding WE, Day R, Holst VA, Wagener MM, Drenning SD, Tweardy DJ: Levels of TGF-alpha and EGFR protein in head and neck squamous cell carcinoma and patient survival. J Natl Cancer Inst 1998;90:824-832.
32 Bentzen SM, Atasoy BM, Daley FM, Dische S, Richman PI, Saunders MI, Trott KR, Wilson GD: Epidermal growth factor receptor expression in pretreatment biopsies from head and neck squamous cell carcinoma as a predictive factor for a benefit from accelerated radiation therapy in a randomized controlled trial. J Clin Oncol 2005;23:5560-5567.
33 Hedback N, Jensen DH, Specht L, Fiehn AM, Therkildsen MH, Friis-Hansen L, Dabelsteen E, von Buchwald C: MiR-21 expression in the tumor stroma of oral squamous cell carcinoma: an independent biomarker of disease free survival. PLoS One 2014;9:e95193.
34 Childs G, Fazzari M, Kung G, Kawachi N, Brandwein-Gensler M, McLemore M, Chen Q, Burk RD, Smith RV, Prystowsky MB, Belbin TJ, Schlecht NF: Low-level expression of microRNAs let-7d and miR-205 are prognostic markers of head and neck squamous cell carcinoma. Am J Pathol 2009;174:736-745.
35 Worku T, Bhattarai D, Ayers D, Wang K, Wang C, Rehman ZU, Talpur HS, Yang L: Long Non-Coding RNAs: the New Horizon of Gene Regulation in Ovarian Cancer. Cell Physiol Biochem 2017;44:948-966.
36 Okugawa Y, Toiyama Y, Hur K, Toden S, Saigusa S, Tanaka K, Inoue Y, Mohri Y, Kusunoki M, Boland CR, Goel A: Metastasis-associated long non-coding RNA drives gastric cancer development and promotes peritoneal metastasis. Carcinogenesis 2014;35:2731-2739.
37 Zheng HT, Shi DB, Wang YW, Li XX, Xu Y, Tripathi P, Gu WL, Cai GX, Cai SJ: High expression of lncRNA MALAT1 suggests a biomarker of poor prognosis in colorectal cancer. Int J Clin Exp Pathol 2014;7:3174-3181.
38 Li L, Geng Y, Feng R, Zhu Q, Miao B, Cao J, Fei S: The Human RNA Surveillance Factor UPF1 Modulates Gastric Cancer Progression by Targeting Long Non-Coding RNA MALAT1. Cell Physiol Biochem 2017;42:2194-2206.
39 Li H, Yuan X, Yan D, Li D, Guan F, Dong Y, Wang H, Liu X, Yang B: Long Non-Coding RNA MALAT1 Decreases the Sensitivity of Resistant Glioblastoma Cell Lines to Temozolomide. Cell Physiol Biochem 2017;42:1192-1201.
40 Xiang Y, Zhang Y, Tang Y, Li Q: MALAT1 Modulates TGF-beta1-Induced Endothelial-to-Mesenchymal Transition through Downregulation of miR-145. Cell Physiol Biochem 2017;42:357-372.
41 Jia LF, Wei SB, Gan YH, Guo Y, Gong K, Mitchelson K, Cheng J, Yu GY: Expression, regulation and roles of miR-26a and MEG3 in tongue squamous cell carcinoma. Int J Cancer 2014;135:2282-2293.
42 Yu JJ, Liu Y, Guo C, Zhang SS, Gong ZJ, Tang YY, Yang LT, He Y, Lian Y, Li XY, Deng H, Liao QJ, Li XL, Li Y, Li GY, Zeng ZY, Xiong W, Yang XM: Upregulated long non-coding RNA LINC00152 expression is associated with progression and poor prognosis of tongue squamous cell carcinoma. J Cancer 2017;8:523-530.
43 Cao W, Liu JN, Liu ZQ, Wang X, Han ZG, Ji T, Chen WT, Zou X: A three-lncRNA signature derived from the Atlas of ncRNA in cancer (TANRIC) database predicts the survival of patients with head and neck squamous cell carcinoma. Oral Oncol 2017;65:94-101.
44 Schwarzenbach H, Hoon DSB, Pantel K: Cell-free nucleic acids as biomarkers in cancer patients. Nat Rev Cancer 2011;11:426-437.
45 Arita T, Ichikawa D, Konishi H, Komatsu S, Shiozaki A, Shoda K, Kawaguchi T, Hirajima S, Nagata H, Kubota T, Fujiwara H, Okamoto K, Otsuji E: Circulating Long Non-coding RNAs in Plasma of Patients with Gastric Cancer. Anticancer Res 2013;33:3185-3193.
46 Isin M, Ozgur E, Cetin G, Erten N, Aktan M, Gezer U, Dalay N: Investigation of circulating lncRNAs in B-cell neoplasms. Clin Chim Acta 2014;431:255-259.
47 Tong YS, Wang XW, Zhou XL, Liu ZH, Yang TX, Shi WH, Xie HW, Lv J, Wu QQ, Cao XF: Identification of the long non-coding RNA POU3F3 in plasma as a novel biomarker for diagnosis of esophageal squamous cell carcinoma. Mol Cancer 2015;14:3.